Numerical Verification of Existence and Inclusion of Solutions for Nonlinear Operator Equations

AUTHOR(S):

Oishi, Shin'ichi; Kashiwagi, Masahide

CITATION:

Oishi, Shin'ichi ...[et al]. Numerical Verification of Existence and Inclusion of Solutions for Nonlinear Operator Equations. 数理解析研究所講究録 1993, 831: 115-128

ISSUE DATE:

1993-04

URL:

http://hdl.handle.net/2433/83368

数理解析研究所講究録
第 831 巻 1993 年 115-128

Numerical Verification of Existence and Inclusion of Solutions for Nonlinear Operator Equations

Shin'ichi Oishi† and Masahide Kashiwagi‡
(大石 進一) (柏木 雅英)

Department of Information and Computer Sciences,
School of Science and Engineering, Waseda University, Tokyo 169, Japan.

e-mail: †oishi@oishi.info.waseda.ac.jp
‡kashi@oishi.info.waseda.ac.jp

Abstract

Abstract nonlinear operator equations of the type

$$f(u) \equiv Lu + Nu = 0,\ u \in D(L)$$

are considered, where L is a densely defined closed linear operator from a Banach space X to an another Banach space Y and N a densely defined nonlinear operator from X to Y. A method is presented for numerical verification and inclusion of solutions for the equations.

1 Introduction

In this paper, we are concerned with abstract nonlinear operator equations of the type

$$f(u) \equiv Lu + Nu = 0,\ u \in D(L) \tag{1}$$

where L is a closed linear operator from a Banach space X to an another Banach space Y, and N a nonlinear operator from X to Y. This type of equations occur in a variety of situations in both pure and applied sciences. Eq. (1) is sometimes called a coincidence equation because one wants to find a point u for which the images under L and $-N$ coincide. The purpose of the paper is to present a method for numerical verification of existence and inclusion of solutions for Eq. (1). That is, in association with a certain approximate solution $\tilde{u}$ of Eq. (1), we present an algorithm which may answer the question as to whether there exists an exact solution u^* in some neighborhood of $\tilde{u}$, and in the affirmative case may give a bound for $u^* - \tilde{u}$. If an error bound for $u^* - \tilde{u}$ can be obtained, we shall say that an inclusion of a solution u^* is obtained. In the following, the domain of the definition of $L, D(L)$, and that for $N, D(N)$, is assumed to be Banach spaces satisfying $D(L) \subset D(N)$. For the sake of simplicity we will denote $D = D(L)$. The norms of D, X, and Y will be denoted by $\|\cdot\|_D$, $\|\cdot\|_X$, and $\|\cdot\|_Y$, respectively. Moreover, the operator norm of a linear continuous operator L_1 from a Banach space X_1 to an another Banach space X_2 is denoted as $\|L_1\|_{L(X_1,X_2)}$.

For the case of $L = d/dt$, in 1965, Urabe[16] has presented a method for numerical verification of existence and inclusion of solutions for Eq. (1) . Then, he[17],[15] and his coresearchers[12],[13], [14] presented numerical verification results of periodic and quasi-periodic solutions for various ordinary differential equations. Urabe's method is based on his convergence theorem of a simplified Newton method for operator equations on suitable function spaces. From the numerical analytic point of view, the crucial point of applying Urabe's convergence theorem is to estimate the operator norm of the inverse of the linearized operator of f. Urabe[16] has also presented a method in which the estimation is derived by obtaining the fundamental matrix of the linearized equation of Eq. (1) through the numerical integration. In 1972, Bouc[1] has shown that this kind of estimation can be accomplished without the numerical integration by using functional analytic techniques. The aim of this paper is to extend Urabe-Bouc's approach. That is, in this paper, we will treat the case in which L is a general closed operator including not only ordinary differential operators but also certain types of partial differential operators such as elliptic operators. Since mathematically rigorous bounds is required in obtaining such an estimate, we have developed a numerical software on which rational arithmetic can be executed. In this system using a continued fraction expansion of rational numbers for the rounding of rational numbers, rounding errors during the numerical estimation are completely taken into account.

Historically, several authors have presented different ways to use computers in proving the existence of solutions for nonlinear operator equations. Kantorovich[5] has presented a convergence theorem of the Newton method on function spaces and treated various kinds of functional equations. Kedem[7] has utilized this Newton-Kantorovich theorem to prove the existence of solutions for certain two-point boundary problems through the numerical estimation. Cesari[2] presented also a method based on the alternative method. Collatz[3] and Schroeder [11] have presented methods based on the monotonicity or the inverse-positivity of the operators. More recently, Kaucher-Miranker[6] presented a method using basies expansions. Nakao[9] has presented an infinite dimensional interval method and treated not only ordinary differential equations but also partial differential equations of various types. Plum[10] has also presented a method based on the eigenvalue estimation. Our method of estimating the operator norm of the linearized operator of f is completely different from these method.

2 Graph Norm Estimate

We consider here the graph norm introduced by L in $D(L)$:

$$\|u\|_L = \|u\|_X + \|Lu\|_Y \quad \text{for } u \in D(L)$$

Since L is closed, $D(L)$ becomes a Banach space with respect to the norm $\|u\|_L$. We denote this Banach space D_L. We assume that N is continuously Fréchet differentiable as a map from D_L to Y. For $u \in D_L$, we assume that the first derivative of N, $DN(u) = S(u)$, can be extended to a bounded linear map from X to Y. In order to verify the existence of solutions for Eq. (1) through the numerical estimation, we introduce now a numerical framework. Let E and F be finite dimensional subspaces of D_L and Y, respectively, with $\dim E = \dim F = m$. Let P and Q be projections from D_L to E and Y to F, respectively. We assume that

$$\|u - Pu\|_X \le c\|Lu\|_Y \quad \text{for } \forall u \in D_L \tag{2}$$

$$QLu = QLPu \quad \text{for } \forall\ u \in D_L \tag{3}$$

and

$$\|Q\|_{L(Y,Y)} \leq 1 \tag{4}$$

hold. Here c is a constant independent of u. It should be noted that for a choice of P we usually suppose that the constant c can be chosen arbitrary small provided that $\dim E$ becomes sufficiently large.

Let $\{e_1, e_2, \cdots, e_m\}$ and $\{v_1, v_2, \cdots, v_m\}$ be bases of E and F, respectively. Then any element $e \in E$ and $v \in F$ can be represented as

$$e = \sum_{n=1}^{m} c_n(e) e_n \tag{5}$$

and

$$v = \sum_{n=1}^{m} d_n(v) v_n, \tag{6}$$

respectively. Here, $c_n(e)$'s and $d_n(v)$'s are suitable linear functionals. Thus maps $A_m : E \to E_m$ and $B_m : F \to F_m$ can be defined as

$$A_m e = (c_1(e), c_2(e), \cdots, c_m(e))^t \tag{7}$$

and

$$B_m v = (d_1(v), d_2(v), \cdots, d_m(v))^t, \tag{8}$$

respectively. Here, the superscript t denotes the transposition of vectors,

$$E_m = \{(c_1(e), c_2(e), \cdots, c_m(e))^t | e \in E\}$$

and

$$F_m = \{(d_1(v), d_2(v), \cdots, d_m(v))^t | v \in F\}.$$

For $\phi = (c_1, c_2, \cdots, c_m)^t \in E_m$ and $d = (d_1, d_2, \cdots, d_m)^t \in F_m$, define

$$\|\phi\|_{E_m} = \| \sum_{n=1}^{m} c_n e_n \|_X \tag{9}$$

and

$$\|d\|_{F_m} = \| \sum_{n=1}^{m} d_n v_n \|_Y. \tag{10}$$

Now, let $\tilde{u} \in E$ be a certain approximate solution of Eq. (1). For example, $\tilde{u}$ is obtained by solving the following determining equation of the Galerkin approximation

$$Q_m f(u) = 0 \text{ for } u \in E_m \tag{11}$$

through the usual floating point arithmetic. Thus $\tilde{u}$ is not an exact solution even for this approximate equation. Then, a linear transformation $J : E_m \to F_m$ can be defined for $\phi = (c_1, c_2, \cdots, c_m)^t \in E_m$ by

$$J\phi = B_m\{Q(L + S(\tilde{u})) \sum_{n=1}^{m} c_n e_n\}. \tag{12}$$

Since E_m and F_m are finite dimensional vector spaces, from now on, J is identified with a matrix. By the definition, we have for $x \in D_L$

$$JA_mPx = B_m\{Q(L + S(\tilde{u}))Px\}. \tag{13}$$

If $\det J \neq 0$, we have

$$A_mPx = J^{-1}B_m\{Q(L + S(\tilde{u}))Px\}, \tag{14}$$

from which we have

$$\begin{aligned} \|Px\|_X &= \|A_mPx\|_{E_m} \\ &\leq \|J^{-1}\|_{L(F_m,E_m)}\|B_mQ(L + S(\tilde{u}))Px\|_{F_m} \\ &\leq M\|Q(L + S(\tilde{u}))Px\|_Y \end{aligned} \tag{15}$$

Here, M is a constant such that

$$\|J^{-1}\|_{L(F_m,E_m)} \leq M. \tag{16}$$

Then, one of our main results can be stated as follows:

Theorem 2.1 Assume that $\det J \neq 0$. Let K and M be constants such that $\|S(\tilde{u})\|_{L(X,Y)} \leq K$ and $\|J^{-1}\|_{L(F_m,E_m)} \leq M$. If $cK(1 + MK) < 1$, then the map $G(\tilde{u}) = L + S(\tilde{u}) : D_L \to Y$ satisfies the following estimate for any $x \in D_L$:

$$\|x\|_L \leq C\|G(\tilde{u})x\|_Y, \tag{17}$$

where

$$C = \frac{(1+c)(1+MK)+M}{1-cK(1+MK)}.$$

□

From this theorem, it is seen that if the constants K and M can be evaluated numerically, then the constant C can be estimated provided $CK(1+MK) < 1$ holds. The rational arithmetic numerical software library has been developed for estimating the constants such as K and M taking the rounding errors of the numerical computation into account. Details will be discussed in later by choosing a suitable example.

It should also be note that Th.2.1 states that the map $G(\tilde{u}) = L + S(\tilde{u}) : D_L \to Y$ is an injection. If this map is also a surjection, it follows that the map has the inverse. Although, this is not the case in general, for the Fredholm operators we can show that the map has the inverse. We recall here the definition of the Fredholm map with an index zero. The continuous linear operator T from a Banach space X_1 to an another Banach space X_2 is call of Fredholm type iff

$$\dim N(T) < \infty$$

and

$$\mathrm{codim} R(T) < \infty.$$

Here, $N(T)$ and $R(T)$ are the null space and the range of the operator T, respectively. codim $R(T)$ is the dimension of the space $X_2/R(T)$. For the Fredholm operator T

$$\mathrm{ind}(T) = \dim N(T) - \mathrm{codim} R(T) \tag{18}$$

is well defied and called the index. If we consider the map $G(\tilde{u})$ is as the map from the Banach space D_L to the another Banach space Y, it becomes continuous.

Corollary 2.1 If $G(\tilde{u})$ is of Fredholm type with the index 0 and if the condtion of Th.2.1 is satisfied, then $G(\tilde{u})$ has the inverse. □

In fact, from Th.2.1 it follow that

$$\dim N(G(\tilde{u})) = 0 \tag{19}$$

which implies codim$R(G(\tilde{u})) = 0$, because the index of $G(\tilde{u})$ is assumed to be zero. Thus it is shown that $G(\tilde{u})$ is also surjective and has the inverse.

Now we define a residual

$$r = \|f(\tilde{u})\|_Y.$$

Let $U_p = B(\tilde{u}, p)$ be the closed ball in D_L centered at $\tilde{u}$ with the radius p. Here, if we assume that $S(u) = DN(u) : D_L \to Y$ is locally Lipschitz continuous:

$$\|S(u) - S(v)\|_{L(D_L,Y)} \leq a_{U_p}\|u - v\|_L \quad \text{for } u, v \in U_p \subset D_L,$$

then we have

Theorem 2.2 Assume that $G(\tilde{u}) : D_L \to Y$ has the inverse and $cK(1 + MK) < 1$ holds. For the sake of simplicity, let $a = a_{U_p}$. If p satisfies

1. $2Cr \leq p$
   and
2. $aCp < 1$,

then there exists a solution u^* of Eq. (1) uniquely in U_p such that

$$\|u^* - \tilde{u}\|_L \leq 2Cr.$$

□

This theorem implies that together with K and M, if the constants r and a can further be estimated numerically, the existence of a solution for Eq. (1) is verified numerically provided that the conditions of Theorem 2.2 are satisfied.

3 Proof of Theorem 2.1

Recall that

$$G(\tilde{u})x = Lx + S(\tilde{u})x, \quad G(\tilde{u}) : D_L \to Y. \tag{20}$$

For $x \in D_L$, we have

$$\begin{aligned} \|x\|_X &\leq \|x - Px\|_X + \|Px\|_X \\ &\leq c\|Lx\|_Y + \|Px\|_X. \end{aligned} \tag{21}$$

From the definition of (20) and (21), it follows

$$\begin{aligned}\|Lx\|_Y &\leq \|G(\tilde{u})x\|_Y + \|S(\tilde{u})x\|_Y \\ &\leq \|G(\tilde{u})x\|_Y + K\|x\|_X \\ &\leq \|G(\tilde{u})x\|_Y + cK\|Lx\|_Y + K\|Px\|_X. \end{aligned} \tag{22}$$

Moreover from (20) and (3), we have

$$QG(\tilde{u})x = QLx + QS(\tilde{u})x = QLPx + QS(\tilde{u})(x - Px + Px).$$

Here, if we put

$$s = QLPx + QS(\tilde{u})Px = Q[G(\tilde{u})x - S(\tilde{u})(x - Px)],$$

using (4) we have

$$\|s\|_Y \leq \|G(\tilde{u})x\|_Y + cK\|Lx\|_Y. \tag{23}$$

Substituting the relation (15)

$$\|Px\|_X \leq M\|s\|_Y \tag{24}$$

and (23) into (22), we have

$$\begin{aligned}\|Lx\|_Y &\leq \|G(\tilde{u})x\|_Y + cK\|Lx\|_Y + MK\|s\|_Y \\ &\leq \|G(\tilde{u})x\|_Y + cK\|Lx\|_Y + MK(\|G(\tilde{u})x\|_Y + Kc\|Lx\|_Y) \\ &= (1 + MK)\|G(\tilde{u})x\|_Y + cK(1 + MK)\|Lx\|_Y.\end{aligned}$$

Thus we have

$$\|Lx\|_Y \leq \frac{1 + MK}{1 - cK(1 + MK)}\|G(\tilde{u})x\|_Y. \tag{25}$$

On the other hand, substituting (24) and (23) into (21), we have

$$\begin{aligned}\|x\|_X &\leq c\|Lx\|_Y + M\|s\|_Y \\ &\leq c\|Lx\|_Y + M(\|G(\tilde{u})x\|_Y + cK\|Lx\|_Y) \\ &= c(1 + MK)\|Lx\|_Y + M\|G(\tilde{u})x\|_Y.\end{aligned}$$

From this and (25), we have

$$\|x\|_X \leq \frac{c(1 + MK) + M}{1 - cK(1 + MK)}\|G(\tilde{u})x\|_Y. \tag{26}$$

Summing up the above-mentioned discussions, we finally have

$$\|x\|_L = \|x\|_X + \|Lx\|_Y \leq \frac{(1 + c)(1 + MK) + M}{1 - cK(1 + MK)}\|G(\tilde{u})x\|_Y$$

provided $cK(1 + MK) < 1$. This proves Theorem 2.1. □

4 Proof of Theorem 2.2

We shall prove Theorem 2.2 by showing that the operator T defined in the below becomes a contraction mapping on U_p under the conditions of Theorem 2.2. Using $G(\tilde{u})^{-1}$, let us define an operator $T : D_L \to D_L$ by

$$Tu = G(\tilde{u})^{-1}(S(\tilde{u})u - Nu).$$

Since $G(\tilde{u})^{-1}$ exists, a fixed point of T is a solution of Eq. (1). In the first place, we shall show that $TU_p \subset U_p$. For any $u \in U_p$, we have

$$\begin{aligned} \|Tu - \tilde{u}\|_L &= \|G(\tilde{u})^{-1}(S(\tilde{u})u - Nu) - \tilde{u}\|_L \\ &= \|G^{-1}(\tilde{u})(S(\tilde{u})u - Nu - G(\tilde{u})\tilde{u})\|_L \\ &\le C\|S(\tilde{u})u - Nu - G(\tilde{u})\tilde{u}\|_Y \\ &= C\|S(\tilde{u})u - Nu - L\tilde{u} - S(\tilde{u})\tilde{u}\|_Y \\ &\le C(\| - Nu + N\tilde{u} - S(\tilde{u})(\tilde{u} - u)\|_Y + r). \end{aligned} \tag{27}$$

Since $L\tilde{u} = f(\tilde{u}) - N\tilde{u}$ and $\|f(\tilde{u})\|_Y = r$. Let

$$R = Nu - N\tilde{u} - S(\tilde{u})(u - \tilde{u}).$$

Using the formula

$$Nu - Nv = \int_0^1 S(u + t(v - u))(v - u)dt,$$

we have an estimate

$$\begin{aligned} \|R\|_Y &= \| \int_0^1 (S(\tilde{u} + t(u - \tilde{u})(u - \tilde{u})) - S(\tilde{u}))(u - \tilde{u})dt\|_Y \\ &= \| \int_0^1 [S(\tilde{u} + t(u - \tilde{u})) - S(\tilde{u})](u - \tilde{u})dt\|_Y \\ &\le a \int_0^1 \|[\tilde{u} + t(u - \tilde{u})] - \tilde{u}\|_Y \|(u - \tilde{u})\|_Y dt \\ &\le \frac{a}{2}\|u - \tilde{u}\|_L^2, \end{aligned} \tag{28}$$

from which, we have

$$\begin{aligned} \|Tu - \tilde{u}\|_L &\le C\left(\frac{a}{2}\|u - \tilde{u}\|_L^2 + r\right) \\ &\le C\left(\frac{a}{2}p^2 + r\right) < p. \end{aligned} \tag{29}$$

This implies $TU_p \subset U_p$.

We now show that T is contractive on U_p. For for $u, v \in U_p$, we have

$$\begin{aligned} \|Tu - Tv\|_L &\le \|G(\tilde{u})^{-1}(S(\tilde{u})u - Nu) - G(\tilde{u})^{-1}(S(\tilde{u})v - Nv)\|_L \\ &= \|G(\tilde{u})^{-1}(S(\tilde{u})(u - v) - (Nu - Nv))\|_L \\ &\le C\|S(\tilde{u})(u - v) - (Nu - Nv)\|_Y \end{aligned}$$

$$
\begin{aligned}
&= C\|\int_0^1 (S(u+t(v-u)) - S(\tilde{u}))(v-u)dt\|_Y \\
&\le C\int_0^1 \|S(u+t(v-u)) - S(\tilde{u})\|_{L(D_L,Y)}\|v-u\|_L dt \\
&\le aCp\|v-u\|_L. \qquad (30)
\end{aligned}
$$

Thus we have

$$\|Tu - Tv\|_L \le aCp\|v-u\|_L. \qquad (31)$$

This shows that T is contractive on U_p. Thus it follows that there exists a unique fixed point u^* of T in U_p. From the relation

$$\|u^* - \tilde{u}\|_L \le \frac{a}{2}Cp\|Tu^* - \tilde{u}\|_L + Cr,$$

we obtain an error bound

$$\|u^* - \tilde{u}\|_L \le 2Cr.$$

This completes the proof. □

5 An Application to An Ordinary Differential Equation

In this section, we study an application of the results in the previous sections to obtain a periodic solution of ordinary differential equations taking the following Duffing equation

$$x'' + Ax' + Bx^3 - C\cos t = 0, t \in J = (0, 2\pi)$$

as an example, where A, B and C are constants. Let $L_2(0,2\pi)$, $H_1(0,2\pi)$ and $H_2(0,2\pi)$ be the Lebesgue space of square integrable functions and the Sobolev spaces with norms

$$\|x\|_2 = \sqrt{\frac{1}{2\pi}\int_0^{2\pi} |x(t)|^2\, dt},$$

$$\|x\|_{H_1} = \sqrt{\|x\|^2 + \|x'\|^2},$$

and

$$\|x\|_{H_2} = \sqrt{\|x\|^2 + \|x'\|^2 + \|x''\|^2},$$

respectively. Let $X = Y = \{x | x \in L_2(0,2\pi) \cap x(t) = -x(t+\pi)\}$. Let us define operators $L : D(L) = X \cap H_2(0,2\pi) \to Y$ and $N : D(L) \to Y$ by

$$Lx = x'' + Ax'$$

and

$$Nx = Bx^3 - C\cos t,$$

respectively. Then, it is well known that L is a closed linear operator from X to Y. Thus the graph norm associated with L is defined as

$$\|x\|_L = \|x\|_2 + \|x'' + Ax'\|_2.$$

For $x \in D(L)$, taking the equation $x(t) = -x(t+\pi)$ we can expand x as

$$x = \sqrt{2}\sum_{n=1}^{\infty}(a_n \cos(2n-1)t + b_n \sin(2n-1)t).$$

Now define a projection operator $P_m : D(L) \to E = P_m D(L)$ by

$$P_m x = \sqrt{2}\sum_{n=1}^{m}(a_n \cos(2n-1)t + b_n \sin(2n-1)t).$$

Then we have

Lemma 5.1

$$\|x - P_m x\|_2 \le \frac{1}{\sqrt{(2m+1)^4 + A^2(2m+1)^2}}\|Lx\|_2$$

for $x \in D(L)$, where $P_m D(L)$ is the image of $D(L)$ by P_m. □

Proof Let

$$x' = \sqrt{2}\sum_{n=1}^{\infty}(a'_n \cos(2n-1)t + b'_n \sin(2n-1)t)$$

and

$$x'' = \sqrt{2}\sum_{n=1}^{\infty}(a''_n \cos(2n-1)t + b''_n \sin(2n-1)t).$$

So we have

$$a'_n = (2n-1)b_n, b'_n = -(2n-1)a_n,$$

and

$$a''_n = -(2n-1)^2 a_n, b''_n = -(2n-1)^2 b_n.$$

Thus if we put

$$x'' + Ax'(t) = \sqrt{2}\sum_{n=1}^{\infty}(\tilde{a}_n \cos(2n-1)t + \tilde{b}_n \sin(2n-1)t),$$

we have

$$\tilde{a}_n = -(2n-1)^2 a_n + (2n-1)Ab_n, \tilde{b}_n = -(2n-1)Aa_n - (2n-1)^2 b_n,$$

or

$$a_n = \frac{-(2n-1)^2\tilde{a}_n - (2n-1)A\tilde{b}_n}{(2n-1)^4 + (2n-1)^2A^2}$$

and

$$b_n = \frac{-(2n-1)^2\tilde{b}_n + (2n-1)A\tilde{a}_n}{(2n-1)^4 + (2n-1)^2A^2}.$$

Let us now consider $\|x - P_m x\|_2^2$. The Perseval equality gives

$$\begin{aligned}
\|x - P_m x\|_2^2 &= \sum_{n=m+1}^{\infty}(a_n^2 + b_n^2) \\
&\le \sum_{n=m+1}^{\infty}\frac{1}{((2n-1)^4 + A^2(2n-1)^2)}(\tilde{a}_n^2 + \tilde{b}_n^2) \\
&\le \frac{1}{(2m+1)^4 + A^2(2m+1)^2}\|Lx\|_2^2.
\end{aligned}$$

Thus we have the desired inequality. □

Moreover, we have

Lemma 5.2 For $x \in H_2(0, 2\pi)$, we have

$$\tilde{b}\|x\|_L \le \|x\|_{H_2} \le b\|x\|_L,$$

where

$$\tilde{b} = \frac{1}{2+A}$$

and

$$b = \sqrt{2(1+A^2)}.$$

□

Proof From the Perseval equality, we have

$$\begin{aligned}
\|x''\|_2^2 &= \sum_{n=1}^{\infty}(a_n''^2 + b_n''^2) \\
&\le \sum_{n=1}^{\infty} \frac{(2n-1)^4((2n-1)^4 + A^2(2n-1)^2)}{((2n-1)^4 + A^2(2n-1)^2)^2}(\tilde{a}_n^2 + \tilde{b}_n^2) \\
&\le (1+A^2)\|Lx\|_2^2,
\end{aligned}$$

and similarly

$$\|x'\|_2^2 \le (1+A^2)\|Lx\|_2^2.$$

These inequalities imply

$$\begin{aligned}
\|x''\|_2^2 + \|x'\|_2^2 + \|x\|_2^2 &\le \|x\|_2^2 + 2(1+A^2)\|Lx\|_2^2 \\
&\le 2(1+A^2)\|x\|_L^2, \qquad (32)
\end{aligned}$$

which is the right half of the desired inequalities.

On the other hand, we have

$$\begin{aligned}
\|x\|_L &= \|x\|_2 + \|x'' + Ax'\|_2 \\
&\le \|x\|_2 + \|x''\|_2 + A\|x'\|_2 \\
&\le (2+A)\|x\|_{H_2}.
\end{aligned}$$

This is the left half of the desired inequality. □

Similarly, we obtain

Lemma 5.3 For $x \in H_2(0, 2\pi)$, we have

$$\|x\|_{H_1} \le \sqrt{1+A^2}\|x\|_L,$$

and

$$\|x'\|_{H_1} \le \sqrt{2(1+A^2)}\|x\|_L.$$

□

We now consider to include 2π-periodic solution of the Duffing equation with $A = 0.1, B = 1$, and $C = 0.4464$. For the purpose, let us consider an approximate equation of Eq. (1) of the following form:

$$P_m f(x) = 0, x \in E = P_m D(L). \tag{33}$$

Here,

$$f(x) = Lx + Nx.$$

Since the so-called determing equation (33) is a finite dimensional equation, its approximate solution can be obtained easily. In fact, the following approximate solution is derived through the Newton method:

$$\begin{aligned}
\tilde{x}(t) &= \frac{12391844444622}{10096283453831}\cos t + \frac{1255301899357}{3264990063609}\sin t \\
&+ \frac{3339800261015}{62230322929326}\cos 3t + \frac{25614353059037}{407715265530912}\sin 3t \\
&+ \frac{30678010753}{50578758054295}\cos 5t + \frac{20268208717}{4200092845578}\sin 5t \\
&- \frac{203050479}{1606019671451}\cos 7t + \frac{19543149859}{75359444598260}\sin 7t \\
&- \frac{9917353}{674649767686}\cos 9t + \frac{27060356}{3079992935547}\sin 9t \\
&- \frac{10029085}{9872509922553}\cos 11t - \frac{80843412}{2002007632142809}\sin 11t \\
&- \frac{353059}{7177837174127}\cos 13t - \frac{925405}{26456112180297}\sin 13t \\
&- \frac{2009793}{1535022779191217}\cos 15t - \frac{1158567}{347492958486574}\sin 15t.
\end{aligned}$$

Now letting $P = Q = P_m$, J is computed through the formula (12). Thanks to the polynomial nonlinearity of the problem, the matrix J can be calculated rigorously. In fact, using the addition formula of the trigonomeric functions and the technique of the automatic differentiation[4], a program for calculating J rigorously can be realized without difficulty. Then, since each element of J is rational number, J^{-1} can be calculated exactly by the rational arithmetic, which is executed on the rational arithmetic library developed by ourselves. Thus, a bound M of $\|J^{-1}\|_{L(F,E)}$ can be evaluated as the Frobenius norm of the matrix J^{-1} free from the rounding errors of numerical computation. Here, $F_m = P_m Y$. Similarly, the residual $\|f(\tilde{x})\|_2$ can be estimated numerically through the Parseval equality free from the numerical computation errors .

In this example, we have an estimate

$$\|S(\tilde{x})\|_{L(X,Y)} \le \|3\tilde{x}^2\|_\infty.$$

Here, $\|x\|_\infty = \max\{|x(t)||0 \le t \le 2\pi\}$. Since

$$\|\sqrt{2}\sum_{n=1}^{m}(a_n \cos(2n-1)t + b_n \sin(2n-1)t)\|_\infty \le \sqrt{2}\sum_{n=1}^{m}\sqrt{|a_n|^2 + |b_n|^2}, \tag{34}$$

the value $K = \|3\tilde{x}^2\|_\infty$ can be evaluated rigorously using the rational arithmetic library. Similarly, $a = a_{B(\tilde{x},p)}$ is estimated as

$$a = \frac{6bd}{\sqrt{2}}(\|\tilde{x}\|_\infty + p),$$

which can also be evaluated rigorously using the rational arithmetic numerical library.

Thus, for the approximate solution $\tilde{x}$, as a result of the estimation, we have

$$M \leq 3.118, r \leq 0.0000000432, K \leq 6.869 \text{ and } p \leq 0.00000474.$$

From these constants, we have

$$C \leq 54.806, a \leq 26.215 \text{ and } aCp \leq 0.00682.$$

Of course, this evaluation is free from the numerical computation errors. For the Duffing equation it is easy to show that the operator $G(\tilde{u})$ becomes a Fredholm operator with the index zero so that the existence of the constant C implies the existence of the inverse of the operator $L + S(\tilde{x})$. Thus it is verified from Theorem 2.2 that, in the ball $\|\tilde{x} - x\|_L \leq 0.00000474$, there exists a locally unique exact solution $x^*$of the Duffing equation. By the Sobolev embedding theorem[8], for $x \in H_1(0, 2\pi)$ we have

$$\|x\|_\infty \leq \sqrt{\frac{2\pi}{\tanh 2\pi}}\|x\|_{H_1},$$

from which we have the following estimate between x^* and $\tilde{x}$ as

$$\begin{aligned} \|\tilde{x} - x^*\|_\infty &\leq d\|\tilde{x} - x^*\|_{H_1} \\ &\leq \frac{bd}{\sqrt{2}}\|\tilde{x} - x^*\|_L \\ &\leq \frac{bdp}{\sqrt{2}}, \\ &\leq 0.0000120, \end{aligned} \tag{35}$$

and

$$\begin{aligned} \|\frac{d\tilde{x}}{dt} - \frac{dx^*}{dt}\|_\infty &\leq d\|\frac{d\tilde{x}}{dt} - \frac{dx^*}{dt}\|_{H_1} \\ &\leq bd\|\tilde{x} - x^*\|_L \\ &\leq 0.0000169, \end{aligned} \tag{36}$$

where $b = \sqrt{1 + A^2}$ and $d = \sqrt{2\pi/\tanh 2\pi}$ so that $bd \leq 3.56261$.

In Fig.1., the outline of the solution is illustrated. In Fig.1 (b), the center line of the three parallel lines indicates $\tilde{x}$ and the other two lines indicate the bound, in which the exact solution x^* is located.

Acknowledgement

The authors would like to express their sincerely thanks to Professor Kazuo Horiuchi for his guidance. Thanks are also due to Professor Takao Kakita for his critical reading of the manuscript and helpful comments.

References

[1] R. Bouc. "Sur la methode de Galerkin-Urabe pour les systemes differentierles periodiques". *Intern. J. Non-Linear Mech.*, 7:175–188, 1972.

[2] L. Cesari. "Functional analysis and periodic solutions of nonlinear equations". *Contributions to differential equations*, 1(2):149–187, 1963.

[3] L. Collatz. *"Functional analysis and numerical mathematics"*. Academic Press, 1966.

[4] M. Iri. "Simultaneous computation of functions, partial derivatives and estimates of rounding errors—complexity and practicality". *Japan J. Applied Mathematics*, 1:223–252, 1984.

[5] L.V. Kantorovich. "Functional analysis and applied mathematics". *Uspeh. Math. Nauk*, 3:89–185, 1948.

[6] E.W. Kaucher and W.L. Miranker. *"Self-validating numerics for function space problems"*. Academic Press, 1984.

[7] G. Kedem. "A posteriori bounds for two-point boundary value problems". *SIAM J. Numer. Anal.*, 18(3):431–448, 1981.

[8] J. T. Marti. "Evaluation of the least constant in Sobolev's inequality for $H^1(0,s)$". *SIAM J. Numer. Anal.*, 20:1239–1242, 1983.

[9] M. Nakao. "A numerical approach to the proof of existence of solutions for elliptic problems ". *Japan J. Appl. Math.*, 5:313–332, 1988.

[10] M. Plum. "Computer-assisted existence proofs for two point boundary value problems". *Computing*, 46:19–34, 1991.

[11] J. Schroeder. "A method for producing verified results for two-point boundary value problems". *Computing Suppl.*, 6:9–22, 1988.

[12] Y. Shinohara. "A geometric method of numerical solutions of nonlinear equations and its application to nonlinear oscillations". *Publ.RIMS, Kyoto Univ.*, 13, 1972.

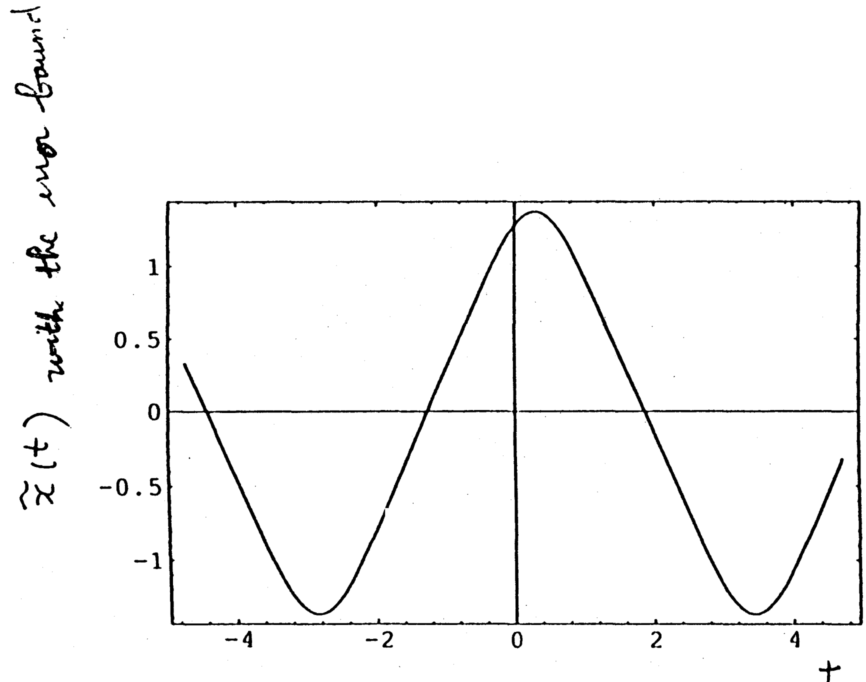
[13] Y. Shinohara. "Numerical analysis of periodic solutions and their periods to autonomous differential systems". *J. Math. Tokushima Univ.*, 11:11–32, 1972.

[14] Y. Shinohara and N. Yamamoto. "Galerkin approximation of periodic solution and its period to van der Pol equation". *J. Math. Tokushima Univ.*, 12:19–42, 1978.

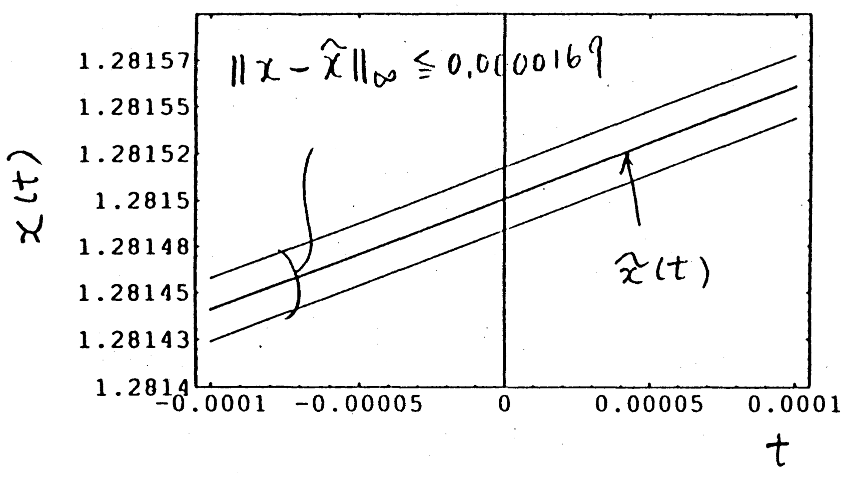
[15] M. Urabe. "Existence theorems of quasiperiodic solutions to nonlinear differential systems". *Funkcialaj Ekvacioj*, 15:75–100, 1972.

[16] M. Urabe. "Galerkin's procedure for nonlinear periodic systems". *Arch. Rational Mech. Anal.*, 20:120–152, 1965.

[17] M. Urabe. "Numerical investigation of subharmonic solution to Duffing's equation". *Publ. RIMS, Kyoto Univ.*, 5:79–112, 1969.



(a) 2π-periodic solution



(b) Scaling of a part of (a).

Fig.1 Result of Inclusion of the 2π-periodic solution of the Duffing Equation